

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
7 March 2002 (07.03.2002)

PCT

(10) International Publication Number
WO 02/18641 A2

(51) International Patent Classification⁷: C12Q 1/68

(21) International Application Number: PCT/IB01/01580

(22) International Filing Date: 30 August 2001 (30.08.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0021286.0 30 August 2000 (30.08.2000) GB

(71) Applicant: GEMINI GENOMICS PLC [GB/GB]; 162
Science Park, Milton Road, Cambridge, Cambridgeshire
CB4 0GH (GB).

(72) Inventors: RISINGER, Carl; Geijersgatan 25C, S-752
56 Uppsala (SE). ANDERSSON, Maria, Kristina; Back-
vagen 14, S-756 52 Uppsala (SE). LEWANDER, Tommy;
Rosenvagen 6, S-756 52 Uppsala (SE). OLAISSON, Erik;
Bonadsvagen 8, S-757 57 Uppsala (SE).

(81) Designated States (*national*): AE, AG, AL, AU, BA, BB,
BG, BR, BZ, CA, CN, CO, CR, CU, CZ, DM, DZ, EC, EE,
GD, GE, HR, HU, ID, IL, IN, IS, JP, KP, KR, LC, LK, LR,
LT, LV, MA, MG, MK, MN, MX, NO, NZ, PH, PL, RO,
SG, SI, SK, TT, UA, UZ, VN, YU, ZA.

(84) Designated States (*regional*): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD,
TG).

Published:

- without international search report and to be republished
upon receipt of that report
- entirely in electronic form (except for this front page) and
available upon request from the International Bureau

For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

(54) Title: DETECTION OF *CYP3A4* AND *CYP2C9* POLYMORPHISMS

(57) Abstract: The invention provides oligonucleotide primer pairs, sequence determination oligonucleotides, and kits for amplifi-
cation and detection of novel single nucleotide polymorphisms in the 5' flanking regions of the *CYP3A4* and *CYP2C9* genes.



WO 02/18641 A2

DETECTION OF *CYP3A4* AND *CYP2C9* POLYMORPHISMS

The present invention is directed to methods of preparing biological samples for nucleic acid analysis using oligonucleotide primers suitable for amplification of the genes encoding the drug-metabolizing cytochrome P450 enzymes *CYP3A4* and *CYP2C9*.
5

BACKGROUND OF THE INVENTION

Xenobiotics are pharmacologically, endocrinologically, or toxicologically active substances foreign to a biological system. Most xenobiotics, including pharmaceutical agents, are metabolized through two successive reactions. Phase I reactions
10 (functionalization reactions), include oxidation, reduction, and hydrolysis, in which a derivatizable group is added to the original molecule. Functionalization prepares the drug for further metabolism in phase II reactions. During phase II reactions (conjugative reactions, which include glucoronidation, sulfation, methylation and acetylation), the functionalized drug is conjugated with a hydrophilic group. The resulting hydrophilic
15 compounds are inactive and excreted in bile or urine. Thus, metabolism can result in detoxification and excretion of the active substance. Alternatively, an inert xenobiotic may be metabolized to an active compound. For example, a pro-drug may be converted to a biologically active therapeutic or toxin.

The cytochrome P450 (CYP) enzymes are involved in the metabolism of many
20 different xenobiotics. CYPs are a superfamily of heme-containing enzymes, found in eukaryotes (both plants and animals) and prokaryotes, and are responsible for Phase I reactions in the metabolic process. In total, over 500 genes belonging to the CYP superfamily have been described and divided into subfamilies, CYP1-CYP27. In humans, more than 35 genes and 7 pseudogenes have been identified. Members of three
25 CYP gene families, CYP1, CYP2, and CYP3, are responsible for the majority of drug metabolism. The human CYPs which are of greatest clinical relevance for the metabolism of drugs and other xenobiotics are CYP1A2, CYP2A6, CYP2C9, CYP2C19, CYP2D6, CYP2E1 and CYP3A4. The liver is the major site of activity of these enzymes, however CYPs are also expressed in other tissues.

The most important drug-metabolizing CYP enzyme is CYP3A4, which is the major CYP expressed in liver. Expression of the gene encoding CYP3A4 (*CYP3A4*) is inducible by many commonly used drugs, such as dexamethasone, rifampicin, and clotrimazole. CYP3A4 is estimated to metabolize more than 60% of all drugs in clinical use, including calcium channel blockers such as nifedipine, immunosuppressants such as cyclosporin A, macrolide antibiotics such as erythromycin, and steroid hormones. In addition, CYP3A4 metabolizes some carcinogens, and may be implicated in an individual's susceptibility to such toxins.

The existence of more than one form of the CYP3A4 enzyme is caused by polymorphisms in the gene which encodes the CYP3A4 enzyme (the gene being denoted in italics, as *CYP3A4*). In fact, almost 20 polymorphisms in the *CYP3A4* gene have been described (see <http://www.imm.ki.se/cypalleles/> for listing). The distribution of particular *CYP3A4* polymorphisms differs among ethnic groups, however, concomitant differences in CYP3A4 activity and responses to drugs which are CYP3A4 substrates remain to be investigated. *CYP3A4*1A* is the wild type gene, corresponding to the cDNA having GenBank Accession No. A18907 and the genomic DNA having GenBank Accession No. AF280107. A number of mutations in the 5' untranslated region of *CYP3A4* have been described. *CYP3A4*1B* is an A to G substitution at position -392. *CYP3A4*1C* is a T to G substitution at position -444. *CYP3A4*1D* is a C to A substitution at position -62. *CYP3A4*1E* is a T to A substitution at position -369. *CYP3A4*1F* is a C to G substitution at -747. The 5' flanking region of *CYP3A4* is set forth in SEQ ID NO:1 and in Figure 1.

WO 01/20025 discloses single nucleotide polymorphisms in various exons, introns, and in the 3' UTR of *CYP3A4*, as well as oligonucleotides for use in diagnosing and treating abnormal expression and/or function of this gene. WO 00/24926 discloses oligonucleotides for use in detecting an A to G point mutation at position -290 of *CYP3A4*. WO 99/13106 discloses polymorphisms in *CYP3A4*, including an A to G substitution at position -392 of the promoter, at the 7th position of the 10 bp NFSE, within oligonucleotides having sequences ACAAGGGCAAGAGAGAGGC (SEQ ID NO:2)

and ACAAGGGCAGGAGAGAGGC (SEQ ID NO:3), with polymorphic variants indicated in bold type.

U.S.Pat.No. 6,174,684 and corresponding WO 00/09752 disclose an A to G variant in the nifedipine-specific regulatory element located at positions -287 to -296 of *CYP3A4*, which is associated with increased risk of prostate cancer and with increased risk of developing leukemia after administration of an epipodophyllotoxin. U.S.Pat.No. 6,174,684 also discloses the oligonucleotides AGGGCAAGAG (SEQ ID NO:4) and AGGGCAGGAG (SEQ ID NO:5), with polymorphic variants indicated in bold type. Rebbeck, *et al.* (1998) *J. Natl. Cancer Inst.* **90**, 1225-1229 also describes this association between prostate cancer, leukemia, and the A to G mutation.

Kuehl, *et al.* (2001) *Nature Genetics* **27**, 383-391 discloses mutations at positions -341, -288, and -43 of the *CYP3A4* promoter, none of which were associated with altered *CYP3A4* activity. Kuehl, *et al.* also discloses differential distribution of these polymorphisms among Caucasians and African Americans.

A second important CYP enzyme is *CYP2C9*, which is active in hydroxylation of such drugs as tolbutamide, phenytoin, S-warfarin, diclofenac, ibuprofen, and losarten. The sequence of *CYP2C9* is set forth in SEQ ID NO:6. Six variants in *CYP2C9* are described on the CYP web site, and another six variant designations are listed without descriptions. The *CYP2C9*1* variant is designated as the wild type. Four of the five polymorphic *CYP2C9* forms described contain mutations in the coding regions of the gene that result in decreased *in vitro* activity, and the remaining variant, *CYP2C9*6*, is a deletion of an A at position 818 which results in a frame shift.

WO00/12757 discloses primer extension assays and kits for detection of the single nucleotide polymorphisms *CYP2C9*2* and *CYP2C9*3*, both of which result in amino acid substitutions.

On the basis of ability of metabolize a marker drug such as nifedipine for *CYP3A4* or S-warfarin for *CYP2C9*, individuals may be characterized as poor metabolizers (PM), intermediate metabolizers (IM), extensive metabolizers (EM) or ultra extensive metabolizers (UEM or UM) for *CYP3A4* or *CYP2C9* substrates, respectively. Poor metabolizers retain the substrate in their bodies for a relatively long period of time,

and are susceptible to toxicity and side effects at “normal” dosages. Ultraextensive metabolizers clear the substrate from their bodies quickly, and require higher than “normal” dosages to achieve a therapeutic effect. Intermediate and extensive metabolizers retain the substrate in their bodies for times between those of PMs and UEMs, and are more likely to respond to “normal” dosages of the drug. However, individuals characterized as IM or EM may differ in drug clearance by as much as 10-fold, and variations in toxicity, side effects, and efficacy for a particular drug may occur among these individuals. However, administration of such drugs to determine an individual’s metabolic capacity may in itself be dangerous, exposing the individual to potential toxic side effects.

A need remains for methods of preparing biological samples that contain the 5' flanking regions of *CYP3A4* or *CYP2C9*, so that this information may be used to predict differential capacities for metabolizing CYP3A4 and CYP2C9 substrates among individuals.

SUMMARY OF THE INVENTION

The present inventors have discovered a novel single nucleotide polymorphism in the 5' flanking region of *CYP3A4*, and six novel polymorphisms in the 5' flanking region of *CYP2C9*. Oligonucleotides have been devised for amplification of the polymorphic regions corresponding to these polymorphisms. These oligonucleotides may be used to prepare biological samples for further analysis of the 5' flanking regions of these genes. The inventors have also devised sequence determination oligonucleotides for use as probes for the novel single nucleotide polymorphisms in *CYP3A4* and *CYP2C9*.

In one embodiment, the invention provides an oligonucleotide primer pair suitable for amplifying a polymorphic region of a 5' flanking region of a *CYP3A4* gene, wherein the polymorphic region corresponds to position 461 of SEQ ID NO:1, which position may also be described as position -644 from the transcription start site of the *CYP3A4* gene.

In another embodiment, the invention provides a sequence determination oligonucleotide for detecting a polymorphic site in a 5' flanking region of a *CYP3A4* gene, said oligonucleotide being complementary to the polymorphic region corresponding to position 461 of SEQ ID NO:1.

In another embodiment, the invention provides a kit for amplification and/or detection of a polymorphic region of the 5' flanking region of a *CYP3A4* gene, said kit comprising at least one oligonucleotide primer pair capable of amplifying the region corresponding to position 461 of SEQ ID NO:1.

In another embodiment, the invention provides an oligonucleotide primer pair suitable for amplifying a polymorphic region of a 5' flanking region of a *CYP2C9* gene, wherein the polymorphic region corresponds to position 957 of SEQ ID NO:6; position 1049 of SEQ ID NO:6; position 1164 of SEQ ID NO:6; position 1526 of SEQ ID NO:6; position 1661 of SEQ ID NO:6; and position 1662 of SEQ ID NO:6. Position 957 of SEQ ID NO:6 may also be described as position -1189 from the transcription start site of the *CYP3C9* gene; position 1049 of SEQ ID NO:6 may also be described as position -1097 from the transcription start site; position 1164 of SEQ ID NO:6 may also be described as position -982 from the transcription start site; position 1526 of SEQ ID

NO:6 may also be described as position -620 from the transcription start site; position 1661 of SEQ ID NO:6 may also be described as position -485 from the transcription start site; and position 1662 of SEQ ID NO:6 may also be described as position -484 from the transcription start site.

5 In yet another embodiment, the invention provides a sequence determination oligonucleotide for detecting a polymorphic site in a 5' flanking region of a *CYP2C9* gene, said oligonucleotide comprising a sequence selected from the group consisting of an oligonucleotide complementary to the polymorphic region corresponding to position 957 of SEQ ID NO:6; an oligonucleotide complementary to the polymorphic region
10 corresponding to position 1049 of SEQ ID NO:6; an oligonucleotide complementary to the polymorphic region corresponding to position 1164 of SEQ ID NO:6; an oligonucleotide complementary to the polymorphic region corresponding to position 1526 of SEQ ID NO:6; an oligonucleotide complementary to the polymorphic region corresponding to position 1661 of SEQ ID NO:6; and an oligonucleotide complementary
15 to the polymorphic region corresponding to position 1662 of SEQ ID NO:6.

 In another embodiment, the invention provides a kit for amplification and/or detection of a polymorphic region corresponding to at least one polymorphic region in the 5' flanking region of the *CYP2C9* gene, said region being selected from the group consisting of position 957 of SEQ ID NO:6; position 1049 of SEQ ID NO:6; position
20 1164 of SEQ ID NO:6; position 1526 of SEQ ID NO:6; position 1661 of SEQ ID NO:6; and position 1662 of SEQ ID NO:6.

BRIEF DESCRIPTION OF THE DRAWINGS

25 Figure 1 shows the sequence of the 5' flanking region of the *CYP3A4* gene as set forth in SEQ ID NO:1, with the novel polymorphic site underlined and highlighted in bold.

 Figure 2 shows the sequence of the 5' flanking region of the *CYP2C9* gene as set forth in SEQ ID NO:6, with the novel polymorphic sites underlined and highlighted in
30 bold.

DETAILED DESCRIPTION OF THE INVENTION

The U.S. patents and publications referenced herein are hereby incorporated by reference.

For the purposes of the invention, certain terms are defined as follows.

5 “Gene” is defined as the genomic sequence of the *CYP2C19* gene.

 “Oligonucleotide” means a nucleic acid molecule preferably comprising from about 8 to about 50 covalently linked nucleotides. More preferably, an oligonucleotide of the invention comprises from about 8 to about 35 nucleotides. Most preferably, an oligonucleotide of the invention comprises from about 10 to about 25 nucleotides. In accordance with the invention, the nucleotides within an oligonucleotide may be analogs or derivatives of naturally occurring nucleotides, so long as oligonucleotides containing such analogs or derivatives retain the ability to hybridize specifically within the polymorphic region containing the targeted polymorphism. Analogs and derivatives of naturally occurring oligonucleotides within the scope of the present invention are exemplified in U.S. Pat. Nos. 4,469,863; 5,536,821; 5,541,306; 5,637,683; 5,637,684; 15 5,700,922; 5,717,083; 5,719,262; 5,739,308; 5,773,601; 5,886,165; 5,929,226; 5,977,296; 6,140,482; WO 00/56746; WO 01/14398, and the like. Methods for synthesizing oligonucleotides comprising such analogs or derivatives are disclosed, for example, in the patent publications cited above and in U.S. Pat. Nos. 5,614,622; 5,739,314; 5,955,599; 20 5,962,674; 6,117,992; in WO 00/75372, and the like. The term “oligonucleotides” as defined herein also includes compounds which comprise the specific oligonucleotides disclosed herein, covalently linked to a second moiety. The second moiety may be an additional nucleotide sequence, for example, a tail sequence such as a polyadenosine tail or an adaptor sequence, for example, the phage M13 universal tail sequence, and the like. 25 Alternatively, the second moiety may be a non-nucleotidic moiety, for example, a moiety which facilitates linkage to a solid support or a label to facilitate detection of the oligonucleotide. Such labels include, without limitation, a radioactive label, a fluorescent label, a chemiluminescent label, a paramagnetic label, and the like. The second moiety may be attached to any position of the specific oligonucleotide, so long as the

oligonucleotide retains its ability to hybridize to the polymorphic regions described herein.

A polymorphic region as defined herein is a portion of a genetic locus that is characterized by at least one polymorphic site. A genetic locus is a location on a chromosome which is associated with a gene, a physical feature, or a phenotypic trait. A polymorphic site is a position within a genetic locus at which at least two alternative sequences have been observed in a population. A polymorphic region as defined herein is said to "correspond to" a polymorphic site, that is, the region may be adjacent to the polymorphic site on the 5' side of the site or on the 3' side of the site, or alternatively may contain the polymorphic site. A polymorphic region includes both the sense and antisense strands of the nucleic acid comprising the polymorphic site, and may have a length of from about 100 to about 5000 base pairs. For example, a polymorphic region may be all or a portion of a regulatory region such as a promoter, 5' UTR, 3' UTR, an intron, an exon, or the like. A polymorphic or allelic variant is a genomic DNA, cDNA, mRNA or polypeptide having a nucleotide or amino acid sequence that comprises a polymorphism. A polymorphism is a sequence variation observed at a polymorphic site, including nucleotide substitutions (single nucleotide polymorphisms or SNPs), insertions, deletions, and microsatellites. Polymorphisms may or may not result in detectable differences in gene expression, protein structure, or protein function. Preferably, a polymorphic region of the present invention has a length of about 1000 base pairs. More preferably, a polymorphic region of the invention has a length of about 500 base pairs. Most preferably, a polymorphic region of the invention has a length of about 200 base pairs.

A haplotype as defined herein is a representation of the combination of polymorphic variants in a defined region within a genetic locus on one of the chromosomes in a chromosome pair. A genotype as used herein is a representation of the polymorphic variants present at a polymorphic site.

The PCR primer pairs of the invention are capable of amplifying the polymorphic region corresponding to position 461 of SEQ ID NO:1, or any of the polymorphic regions corresponding to position 957 of SEQ ID NO:6; position 1049 of SEQ ID NO:6; position

1164 of SEQ ID NO:6; position 1526 of SEQ ID NO:6; position 1661 of SEQ ID NO:6; and position 1662 of SEQ ID NO:6. Specific oligonucleotide primer pairs of the invention, for amplifying position 461 of SEQ ID NO:1, may comprise sequences selected from the group consisting of SEQ ID NO:7 and SEQ ID NO:8; and SEQ ID NO:9 and SEQ ID NO:10. For amplifying only position 957 of SEQ ID NO:6, an oligonucleotide primer pair comprising the sequences set forth in SEQ ID NO:19 and SEQ ID NO:20 may be used. Alternatively, positions 957 and 1049 of SEQ ID NO:6 may be amplified using an oligonucleotide primer pair comprising the sequences set forth in SEQ ID NO:21 and SEQ ID NO:22; or positions 957, 1049, and 1164 may be amplified using an oligonucleotide primer pair comprising the sequences set forth in SEQ ID NO:23 and SEQ ID NO:24. Position 1164 of SEQ ID NO:6 may also be amplified using an oligonucleotide primer pair comprising the sequences set forth in SEQ ID NO:25 and SEQ ID NO:26. Positions 1526, 1661, and 1662 of SEQ ID NO:6 may be amplified using an oligonucleotide primer pair comprising the sequences set forth in SEQ ID NO:27 and SEQ ID NO:28. Positions 1661 and 1662 of SEQ ID NO:6 may be amplified using an oligonucleotide primer pair selected from the group consisting of an oligonucleotide primer pair comprising the sequences set forth in SEQ ID NO:29 and SEQ ID NO:30 and an oligonucleotide primer pair comprising the sequences set forth in SEQ ID NO:31 and SEQ ID NO:32.

Each of the PCR primer pairs of the invention may be used in any PCR method. For example, a PCR primer pair of the invention may be used in the methods disclosed in U.S.Pat.Nos. 4,683,195; 4,683,202, 4,965,188; 5,656,493; 5,998,143; 6,140,054; WO 01/27327; WO 01/27329; and the like. The PCR pairs of the invention may also be used in any of the commercially available machines that perform PCR, such as any of the GeneAmp[®] Systems available from Applied Biosystems.

The oligonucleotides of the invention may be used to determine the sequence of the polymorphic regions of SEQ ID NO:1 or SEQ ID NO:6 as defined herein. In one embodiment, an oligonucleotide of the invention comprises a sequence selected from the group consisting of SEQ ID NO:11; SEQ ID NO:12; SEQ ID NO:13; SEQ ID NO:14; SEQ ID NO:15; SEQ ID NO:16; SEQ ID NO:17; and SEQ ID NO:18, for determining

the sequence of the novel polymorphic region of *CYP3A4* corresponding to position 461 of SEQ ID NO:1. In another embodiment, for determining the sequence of the polymorphic region of *CYP2C9* corresponding to position 957 of SEQ ID NO:6, an oligonucleotide of the invention comprises a sequence selected from the group consisting of SEQ ID NO:33; SEQ ID NO:34; SEQ ID NO:43; SEQ ID NO:44; SEQ ID NO:53; SEQ ID NO:58; SEQ ID NO:63; and SEQ ID NO:68. In another embodiment, for determining the sequence of the polymorphic region of *CYP2C9* corresponding to position 1049 of SEQ ID NO:6, an oligonucleotide of the invention comprises a sequence selected from the group consisting of SEQ ID NO:35; SEQ ID NO:36; SEQ ID NO:45; SEQ ID NO:46; SEQ ID NO:54; SEQ ID NO:59; SEQ ID NO:64; and SEQ ID NO:69. In another embodiment, for determining the sequence of the polymorphic region of *CYP2C9* corresponding to position 1164 of SEQ ID NO:6, an oligonucleotide of the invention comprises a sequence selected from the group consisting of SEQ ID NO:37; SEQ ID NO:38; SEQ ID NO:45; SEQ ID NO:48; SEQ ID NO:55; SEQ ID NO:60; SEQ ID NO:65; and SEQ ID NO:70. In another embodiment, for determining the sequence of the polymorphic region of *CYP2C9* corresponding to position 1526 of SEQ ID NO:6, an oligonucleotide of the invention comprises a sequence selected from the group consisting of SEQ ID NO:39; SEQ ID NO:40; SEQ ID NO:49; SEQ ID NO:50; SEQ ID NO:56; SEQ ID NO:61; SEQ ID NO:66; and SEQ ID NO:71. In another embodiment, for determining the sequences of the polymorphic region of *CYP2C9* corresponding to either of positions 1661 or 1662 of SEQ ID NO:6, an oligonucleotide of the invention comprises a sequence selected from the group consisting of SEQ ID NO:41; SEQ ID NO:42; SEQ ID NO:51; SEQ ID NO:52; SEQ ID NO:57; SEQ ID NO:62; SEQ ID NO:67; and SEQ ID NO:72.

Those of ordinary skill will recognize that oligonucleotides complementary to the polymorphic regions described herein must be capable of hybridizing to the polymorphic regions under conditions of stringency such as those employed in primer extension-based sequence determination methods, restriction site analysis, nucleic acid amplification methods, ligase-based sequencing methods, methods based on enzymatic detection of mismatches, microarray-based sequence determination methods, and the like. The

oligonucleotides of the invention may be synthesized using known methods and machines, such as the ABI™3900 High Throughput DNA Synthesizer and the Expedite™ 8909 Nucleic Acid Synthesizer, both of which are available from Applied Biosystems (Foster City,CA).

5 The oligonucleotides of the invention may be used, without limitation, as *in situ* hybridization probes or as components of diagnostic assays. Numerous oligonucleotide-based diagnostic assays are known. For example, primer extension-based nucleic acid sequence detection methods are disclosed in U.S.Pat.Nos. 4,656,127; 4,851,331; 5,679,524; 5,834,189; 5,876,934; 5,908,755; 5,912,118; 5,976,802; 5,981,186; 6,004,744; 10 6,013,431; 6,017,702; 6,046,005; 6,087,095; 6,210,891; WO 01/20039; and the like. Primer extension-based nucleic acid sequence detection methods using mass spectrometry are described in U.S.Pat.Nos. 5,547,835; 5,605,798; 5,691,141; 5,849,542; 5,869,242; 5,928,906; 6,043,031; 6,194,144, and the like. The oligonucleotides of the invention are also suitable for use in ligase-based sequence determination methods such as those disclosed in U.S.Pat.Nos. 5,679,524 and 5,952,174, WO 01/27326, and the like. 15 The oligonucleotides of the invention may be used as probes in sequence determination methods based on mismatches, such as the methods described in U.S.Pat.Nos. 5,851,770; 5,958,692; 6,110,684; 6,183,958; and the like. In addition, the oligonucleotides of the invention may be used in hybridization-based diagnostic assays such as those described 20 in U.S.Pat.Nos. 5,891,625; 6,013,499; and the like.

 The oligonucleotides of the invention may also be used as components of a diagnostic microarray. Methods of making and using oligonucleotide microarrays suitable for diagnostic use are disclosed in U.S.Pat.Nos. 5,492,806; 5,525,464; 5,589,330; 5,695,940; 5,849,483; 6,018,041; 6,045,996; 6,136,541; 6,142,681; 6,156,501; 6,197,506; 25 6,223,127; 6,225,625; 6,229,911; 6,239,273; WO 00/52625; WO 01/25485; WO 01/29259; and the like.

 The invention is also embodied in a kit comprising at least one oligonucleotide primer pair of the invention. When the kit is used for amplification and detection of *CYP3A4* polymorphisms, it will comprise an oligonucleotide primer pair suitable for 30 amplification of the polymorphic region corresponding to position 461 of SEQ ID NO:1.

Specific primer pairs in this embodiment are selected from the group consisting of SEQ ID NO:7 and SEQ ID NO:8; and SEQ ID NO:9 and SEQ ID NO:10. This embodiment of the kit of the invention may optionally comprise a sequence determination oligonucleotide selected from the group consisting of SEQ ID NO:11; SEQ ID NO:12; SEQ ID NO:13; SEQ ID NO:14; SEQ ID NO:15; SEQ ID NO:16; SEQ ID NO:17; and SEQ ID NO:18.

When the kit of the invention is used for amplification and detection of polymorphisms in the 5' flanking region of *CYP2C9*, it will comprise at least one oligonucleotide primer pair, wherein the primer pair is capable of amplifying a polymorphic region selected from the group consisting of the polymorphic region corresponding to position 957 of SEQ ID NO:6; the polymorphic region corresponding to position 1049 of SEQ ID NO:6; the polymorphic region corresponding to position 1164 of SEQ ID NO:6; the polymorphic region corresponding to position 1526 of SEQ ID NO:6; the polymorphic region corresponding to position 1661 of SEQ ID NO:6; and the polymorphic region corresponding to position 1662 of SEQ ID NO:6. This embodiment may optionally further comprise a sequence determination oligonucleotide for detecting a polymorphic variant at any or all of the polymorphic sites corresponding to positions 957, 1049, 1164, 1526, 1661 and 1662 of SEQ ID NO:6.

The kit of the invention may also comprise a polymerizing agent, for example, a thermostable nucleic acid polymerase such as those disclosed in U.S.Pat.Nos. 4,889,818; 6,077,664, and the like. The kit of the invention may also comprise chain elongating nucleotides, such as dATP, dTTP, dGTP, dCTP, and dITP, including analogs of dATP, dTTP, dGTP, dCTP and dITP, so long as such analogs are substrates for a thermostable nucleic acid polymerase and can be incorporated into a growing nucleic acid chain. The kit of the invention may also include chain terminating nucleotides such as ddATP, ddTTP, ddGTP, ddCTP, and the like. In a preferred embodiment, the kit of the invention comprises at least two oligonucleotide primer pairs, a polymerizing agent, chain elongating nucleotides, at least two sequence determination oligonucleotides and at least one chain terminating nucleotide. The kit of the invention may optionally include buffers, vials, microtiter plates, and instructions for use.

The examples set forth below are provided as illustration and are not intended to limit the scope and spirit of the invention as specifically embodied therein.

EXAMPLE 1

IDENTIFICATION OF *CYP3A4* POLYMORPHISM

The study was performed in accordance with the principles stated in the Declaration of Helsinki as reviewed in Tokyo 1975 and Venice 1983, Hong Kong 1989 and Somerset West 1996. Ten samples (Swedish Caucasians) were selected and used for identification of polymorphisms in the 5' flanking region of *CYP3A4*.

White blood cells isolated from a blood sample drawn from the brachial vein serve as the source of the genomic DNA for the analyses. The DNA was extracted by guanidine thiocyanate method or QIAamp Blood Kit (QIAGEN, Venlo, The Netherlands). The genes included in the study were amplified by PCR and the DNA sequences were determined by full sequencing. All genetic analyses were performed according to Good Laboratory Practice and Standard Operating Procedures. Case Report Forms were designed and used for clinical and genetic data collection. Data was entered and stored in a relational database at Gemini Genomics AB, Uppsala. To secure consistency between the Case Report Forms and the database, data was checked either by double data entry or proofreading. After a Clean File was declared the database was protected against changes. By using the program Stat/TransferTM the database was transferred to SAS data sets. The SASTM system was used for tabulations and statistical evaluations. Genotypes were also correlated against the metabolic ratio.

PCR-fragments were amplified with TaqGOLD polymerase (Applied Biosystems) using Robocycler (Stratagene) or GeneAmp PCR system 9700 (Applied Biosystems). Preferentially, the amplified fragments were 300-400 bp, and the region to be read did not exceed 300 bp. PCR reactions were carried out according to the basic protocol set forth in Table 1, with modifications as indicated in Table 2 for specific primer pairs, which are shown in Table 3. For the GeneAmp PCR 9700 machine the profile used was 10 minutes at 95°, 40 x (45 seconds at 90°, 45 seconds at 60°, 45 seconds at 72°), 5 minutes at 72° and 22° until removed.

Table 1

Solution	Stock Concentration	PCR (μl)
H ₂ O	.	33.2
PCR buffer	10x	5.0
MgCl ₂	25 mM	2.0
dNTP	2.5 mM	2.5
primer 1	10μM	1.0
primer 2	10μM	1.0
Taq-gold polymerase	5 μ/μl	0.3
DNA samples	2 ng/μl	5.0
TOTAL		50.0

Table 2

SEQ ID NO:s	Polymorphic Site	Modification from basic protocol (Table 1)	Detection method
7, 8	461	62° annealing temperature	Full sequencing
9, 10	461	3 μl MgCl ₂ , 58° annealing temperature, 50 cycles	Full sequencing

5

Table 3

Polymorphic Site	Primer Pair
461	SEQ ID NO:7 CCAGCCTGAAAGTGCAGAGA SEQ ID NO:8 TCTTAGAGTCTTTCCTCACCAAAC
461	SEQ ID NO:9 CATGCCCTGTCTCTCCTTTA SEQ ID NO:10 CCATCCCCTTCATGCAATC

10 The optimized condition specified in Table 2 were required to distinguish *CYP3A4* from the closely related gene-family members *CYP3A5*, and *CYP3A7*. Use of the basic protocol will lead to problems when amplifying *CYP3A4*-specific amplicons of 300-400 bp

containing the polymorphisms of interest, unless a nested PCR approach is carried out. The nested PCR approach was not used because of the high risk of contamination when using a nested PCR approach and the high risk of typing errors as a consequence. The modifications shown in Table 2 were optimized and reaction parameters were balanced in such a way that nested PCR was avoided.

For full sequencing, one of the PCR-primers in a primer pair was designed for sequencing by addition of a 29 nucleotide tail complementary to M13 at its 5'-end, namely the nucleotides AGTCACGACGTTGTAAAACGACGGCCAGT. Thus, the entire PCR-product was sequenced from the tailed PCR-primer.

The additional oligonucleotides set forth in Tables 4 through 7 were identified as being suitable for detection of the SNP at positions 461 of the 5' flanking region of the *CYP3A4* gene as depicted in SEQ ID NO:1.

Table 4 sets forth oligonucleotides representing the coding (sense) strand complementary to the polymorphic region corresponding to the novel polymorphism found in the study population. The underlined letter indicates polymorphic position in the sequence context. All sequences are shown in 5' to 3' direction.

Table 4

Polymorphic Site	Sequence	Note
461	SEQ ID NO:11: AGCAC <u>C</u> CTGGT SEQ ID NO:12: AGCAC <u>G</u> CTGGT	C variant G variant

Table 5 sets forth oligonucleotides representing the non-coding (anti-sense) strand complementary to the polymorphic region corresponding to the novel polymorphism found in the study population. The underlined letter indicates polymorphic position in the sequence context. All sequences are shown in 5' to 3' direction.

Table 5

Polymorphic Site	Sequence	Note
461	SEQ ID NO:13: ACCAG <u>G</u> GTGCT SEQ ID NO:14: ACCAG <u>C</u> GTGCT	Antisense G variant Antisense C variant

The sequences of Table 6 represent the 5'-sequence to the novel polymorphic site on the coding (sense) strand (SEQ ID NO: 15) and non-coding (anti-sense) strand (SEQ ID NO:s 16). All sequences are shown in 5' to 3' direction.

5

Table 6

Polymorphic Site	Sequence	Note
461	SEQ ID NO:15: GTGTGTACAGC SEQ ID NO:16: GCTGTACACAC	Sense 5' Antisense 5'

The sequences of Table 7 represent the 3'-sequence to the novel polymorphic site on the non-coding (anti-sense) strand (SEQ ID NO:17) and the coding (sense) strand (SEQ ID NO:18). All sequences are shown in 5' to 3' direction.

10

Table 9

Polymorphic Site	Sequence	Note
461	SEQ ID NO:17: TGGTCCCTACC SEQ ID NO:18: GGTAGGGACCA	Antisense 3' Sense 3'

15

EXAMPLE 2

IDENTIFICATION OF *CYP2C9* POLYMORPHISMS

The study was performed in accordance with the principles stated in the Declaration of Helsinki as reviewed in Tokyo 1975 and Venice 1983, Hong Kong 1989 and Somerset West 1996. Ten samples (Swedish Caucasians) were selected and used for identification of polymorphisms in the 5' flanking region of *CYP2C9*.

White blood cells isolated from a blood sample drawn from the brachial vein serve as the source of the genomic DNA for the analyses. The DNA is extracted by guanidine thiocyanate method or QIAamp Blood Kit (QIAGEN, Venlo, The Netherlands). The genes included in the study were amplified by PCR and the DNA sequences were determined by full sequencing. All genetic analyses were performed according to Good Laboratory Practice and Standard Operating Procedures. Case Report Forms were designed and used for clinical and genetic data collection. Data was entered and stored in a relational database at Gemini Genomics AB, Uppsala. To secure consistency between the Case Report Forms and the database, data was checked either by double data entry or proofreading. After a Clean File was declared the database was protected against changes. By using the program Stat/TransferTM the database was transferred to SAS data sets. The SASTM system was used for tabulations and statistical evaluations. Genotypes were also correlated against the metabolic ratio.

PCR-fragments were amplified with TaqGOLD polymerase (Applied Biosystems) using Robocycler (Stratagene) or GeneAmp PCR system 9700 (Applied Biosystems). Preferentially, the amplified fragments were 300-400 bp, and the region to be read did not exceed 300 bp. PCR reactions were carried out according to the basic protocol set forth in Table 10, with modifications as indicated in Table 11 for specific primer pairs, which are shown in Table 12. For the GeneAmp PCR 9700 machine the profile used was 10 minutes at 95°, 40 x (45 seconds at 90°, 45 seconds at 60°, 45 seconds at 72°), 5 minutes at 72° and 22° until removed.

Table 10

Solution	Stock Concentration	PCR (μl)
H ₂ O		33.2
PCR buffer	10x	5.0
MgCl ₂	25 mM	2.0
dNTP	2.5 mM	2.5
primer 1	10μM	1.0
primer 2	10μM	1.0
Taq-gold polymerase	5 μ/μl	0.3
DNA samples	2 ng/μl	5.0
TOTAL		50.0

Table 11

SEQ ID NO:s	Polymorphic Site	Modification from basic protocol (Table 10)	Detection method
19, 20	957	58° annealing temperature	Full sequencing
21, 22	957 & 1049	3 μl MgCl ₂ , 62° annealing temperature	Full sequencing
23, 24	957, 1049 & 1164	58° annealing temperature	Full sequencing
25, 26	1164	3 μl MgCl ₂ , 62° annealing temperature , 50 cycles	Full sequencing
27, 28	1526, 1661 & 1662		Full sequencing
29, 30	1661 & 1662	3 μl MgCl ₂ , 62° annealing temperature , 50 cycles	Full sequencing
31, 32	1661 & 1662		Full sequencing

5

10

Table 12

Polymorphic Site	Primer Pair	
957	SEQ ID NO:19	CACTAGGGAATTTAGAACAAATATG
	SEQ ID NO:20	GCACAGAAAGCAAAGGAAATTAT
957 & 1049	SEQ ID NO:21	TGTATTTAGATCCTCAACTCAGTATGT
	SEQ ID NO:22	GGATCTCCCTTCTCCATCACT
957, 1049 & 1164	SEQ ID NO:23	GGTCCATTTAGTGATTTCCTAC
	SEQ ID NO:24	ATACACCACATTTATTCTGTTCATA
1164	SEQ ID NO:25	CCAAATTTTCCCTCAGTTACA
	SEQ ID NO:26	TTGGTGCCACACAGCTCATA
1526, 1661 & 1662	SEQ ID NO:27	GCCTTCAGGAATTTTTTTTA
	SEQ ID NO:28	CCAGTTGGGAATATATGATTTAACA
1661 & 1662	SEQ ID NO:29	GCTGCTGTATTTTAGTAGGCTATA
	SEQ ID NO:30	CGTTCCATTGTCCACTCTGTAC
1661 & 1662	SEQ ID NO:31	TCAAGGCAGCTCTGGTGTA
	SEQ ID NO:32	AGTTGGGAATATATGATTTAACAGA

The optimized condition specified in Table 11 were required to distinguish *CYP2C9* from the closely related gene-family members *CYP2C8*, *CYP2C18* and *CYP2C19*. Use of the basic protocol will lead to problems when amplifying *CYP2C9*-specific amplicons of 300-400 bp containing the polymorphisms of interest, unless a nested PCR approach is carried out. The nested PCR approach was not used because of the high risk of contamination when using a nested PCR approach and the high risk of typing errors as a consequence. The modifications shown in Table 11 were optimized and reaction parameters were balanced in such a way that nested PCR was avoided.

For full sequencing, one of the PCR-primers in a primer pair was designed for sequencing by addition of a 29 nucleotide tail complementary to M13 at its 5'-end, namely the nucleotides AGTCACGACGTTGTAAAACGACGGCCAGT. Thus, the entire PCR-product was sequenced from the tailed PCR-primer.

The additional oligonucleotides set forth in Tables 13 through 16 were identified as being suitable for detection of the SNPs at positions 957, 1049, 1164, 1526, 1661 and/or 1662 of the 5' flanking region of the *CYP2C9* gene as depicted in SEQ ID NO:6.

Table 13 sets forth oligonucleotides representing the coding (sense) strand complementary to the polymorphic region corresponding to the polymorphisms found in the study population. The underlined letter indicates polymorphic position in the sequence context. All sequences are shown in 5' to 3' direction.

5

Table 13

Polymorphic Site	Sequence	Note
957	SEQ ID NO:33: ATCTT <u>C</u> TATTG SEQ ID NO:34: ATCTT <u>T</u> TATTG	C variant T variant
1049	SEQ ID NO:35: ACAAT <u>A</u> GAAAG SEQ ID NO:36: ACAAT <u>G</u> GAAAG	A variant G variant
1164	SEQ ID NO:37: ATGG <u>A</u> AGGG SEQ ID NO:38: ATGG <u>A</u> AAGGG	G variant A variant
1526	SEQ ID NO:39: TTAAT <u>G</u> GTAAA SEQ ID NO:40: TTAAT <u>T</u> GTAAA	G variant T variant
1661 & 1662	SEQ ID NO:41: GGATT <u>T</u> CATTAT SEQ ID NO:42: GGATT <u>A</u> AATTAT	TC variants AA variants

Table 14 sets forth oligonucleotides representing the non-coding (anti-sense) strand complementary to the polymorphic region corresponding to the polymorphisms found in the study population. The underlined letter indicates polymorphic position in the sequence context. All sequences are shown in 5' to 3' direction.

10

Table 14

Polymorphic Site	Sequence	Note
957	SEQ ID NO:43: CAATAG <u>A</u> AAGAT SEQ ID NO:44: CAATA <u>A</u> AAGAT	Antisense G variant Antisense A variant
1049	SEQ ID NO:45: CTTTCTATTGT SEQ ID NO:46: CTTTCCATTGT	Antisense T variant Antisense C variant
1164	SEQ ID NO:47: CCCTT <u>C</u> TCCAT SEQ ID NO:48: CCCTT <u>T</u> TCCAT	Antisense C variant Antisense T variant
1526	SEQ ID NO:49: TTTAC <u>C</u> ATTAA SEQ ID NO:50: TTTACAATTAA	Antisense C variant Antisense A variant
1661 & 1662	SEQ ID NO:51: ATAATG <u>A</u> AATCC SEQ ID NO:52: ATAATT <u>T</u> AATCC	Antisense GA variants Antisense TT variant

15

The sequences of Table 15 represent the 5'-sequence to the polymorphic sites on the coding (sense) strand (SEQ ID NO:s 53-57) and non-coding (anti-sense) strand (SEQ ID NO:s 58-67). All sequences are shown in 5' to 3' direction.

Table 15

Polymorphic Site	Sequence	Note
957	SEQ ID NO:53: TACCTCCCATC SEQ ID NO:58: GATGGGAGGTA	Sense 5' Antisense 5'
1049	SEQ ID NO:54: AACCAAAAACA SEQ ID NO:59: TGTTTTTGGTT	Sense 5' Antisense 5'
1164	SEQ ID NO:55: CTGCAGTGATG SEQ ID NO:60: CATCACTGCAG	Sense 5' Antisense 5'
1526	SEQ ID NO:56: TAGGGGGTTTA SEQ ID NO:61: TAAACCCCTA	Sense 5' Antisense 5'
1661 & 1662	SEQ ID NO:57: ATTTGAAAAGGA SEQ ID NO:62: TCCTTTCAAAT	Sense 5' Antisense 5'

- 5 The sequences of Table 16 represent the 3'-sequence to the polymorphic sites on the non-coding (anti-sense) strand (SEQ ID NO:s 63-67) and the coding (sense) strand (SEQ ID NO:s 68-72). All sequences are shown in 5' to 3' direction.

Table 16

Polymorphic Site	Sequence	Note
957	SEQ ID NO:63: TGTGGATGCAA SEQ ID NO:68: TTGCATCCACA	Antisense 3' Sense 3'
1049	SEQ ID NO:64: CATGGCTGCTT SEQ ID NO:69: AAGCAGCCATG	Antisense 3' Sense 3'
1164	SEQ ID NO:65: AGGGATCTCCC SEQ ID NO:70: GGGAGATCCCT	Antisense 3' Sense 3'
1526	SEQ ID NO:66: TAAACACCTTT SEQ ID NO:71: AAAGGTGTTTA	Antisense 3' Sense 3'
1661 & 1662	SEQ ID NO:67: TGTTCTTTATA SEQ ID NO:72: TATAAAGAACA	Antisense 3' Sense 3'

10

CLAIMS

1. An oligonucleotide primer pair suitable for amplifying a polymorphic region of a 5' flanking region of a *CYP3A4* gene, wherein the polymorphic region corresponds to position 816 of SEQ ID NO:1.

5

2. The primer pair of claim 1, having sequences selected from the group consisting of SEQ ID NO: 7 and SEQ ID NO:8 and SEQ ID NO:9 and SEQ ID NO:10.

3. A sequence determination oligonucleotide for detecting a polymorphic site in a 5' flanking region of a *CYP3A4* gene, said oligonucleotide being complementary to the polymorphic region corresponding to position 461 of SEQ ID NO:1.

4. The oligonucleotide of claim 3, comprising a sequence selected from the group consisting of SEQ ID NO:11; SEQ ID NO:12; SEQ ID NO:13; SEQ ID NO:14; SEQ ID NO:15; SEQ ID NO:16; SEQ ID NO:17; and SEQ ID NO:18.

5. A kit comprising at least one oligonucleotide primer pair capable of amplifying the region corresponding to position 461 of SEQ ID NO:1.

6. The kit of claim 5, wherein the primer pair comprises sequences selected from the group consisting of SEQ ID NO: 7 and SEQ ID NO:8 and SEQ ID NO:9 and SEQ ID NO:10.

7. The kit of claim 5, further comprising a sequence determination oligonucleotide complementary to the polymorphic region corresponding to position 461 of SEQ ID NO:1.

8. The kit of claim 7, wherein the oligonucleotide comprises a sequence selected from the group consisting of SEQ ID NO:11; SEQ ID NO:12; SEQ ID NO:13; SEQ ID NO:14; SEQ ID NO:15; SEQ ID NO:16; SEQ ID NO:17; and SEQ ID NO:18.

9. An oligonucleotide primer pair suitable for amplifying a polymorphic region of a 5' flanking region of a *CYP2C9* gene, wherein the polymorphic region corresponds to position 957 of SEQ ID NO:6; position 1049 of SEQ ID NO:6; position 1164 of SEQ ID NO:6; position 1526 of SEQ ID NO:6; position 1661 of SEQ ID NO:6; and position 1662 of SEQ ID NO:6.

10. The primer pair of claim 9, having a sequence selected from the group consisting of SEQ ID NO: 19 and SEQ ID NO:20; SEQ ID NO:21 and SEQ ID NO:22; SEQ ID NO:23 and SEQ ID NO:24; SEQ ID NO:25 and SEQ ID NO:26; SEQ ID NO:27 and SEQ ID NO:28; SEQ ID NO:29 and SEQ ID NO:30; and SEQ ID NO:31 and SEQ ID NO:32.

11. A sequence determination oligonucleotide for detecting a polymorphic site in a 5' flanking region of a *CYP2C9* gene, said oligonucleotide comprising a sequence selected from the group consisting of an oligonucleotide complementary to the polymorphic region corresponding to position 957 of SEQ ID NO:6; an oligonucleotide complementary to the polymorphic region corresponding to position 1049 of SEQ ID NO:6; an oligonucleotide complementary to the polymorphic region corresponding to position 1164 of SEQ ID NO:6; an oligonucleotide complementary to the polymorphic region corresponding to position 1526 of SEQ ID NO:6; an oligonucleotide complementary to the polymorphic region corresponding to position 1661 of SEQ ID NO:6; and an oligonucleotide complementary to the polymorphic region corresponding to position 1662 of SEQ ID NO:6.

12. The oligonucleotide of claim 11, comprising a sequence selected from the group consisting of SEQ ID NO:33; SEQ ID NO:34; SEQ ID NO:35; SEQ ID NO:36; SEQ ID NO:37; SEQ ID NO:38; SEQ ID NO:39; SEQ ID NO:40; SEQ ID NO:41; SEQ ID NO:42; SEQ ID NO:43; SEQ ID NO:44; SEQ ID NO:45; SEQ ID NO:46; SEQ ID NO:47; SEQ ID NO:48; SEQ ID NO:49; SEQ ID NO:50; SEQ ID NO:51; SEQ ID NO:52; SEQ ID NO:53; SEQ ID NO:54; SEQ ID NO:55; SEQ ID NO:56; SEQ ID

NO:57; SEQ ID NO:58; SEQ ID NO:59; SEQ ID NO:60; SEQ ID NO:61; SEQ ID NO:62; SEQ ID NO:63; SEQ ID NO:64; SEQ ID NO:65; SEQ ID NO:66; SEQ ID NO:67 and SEQ ID NO:68.

- 5 13. A kit comprising at least one oligonucleotide primer pair, wherein the primer pair is capable of amplifying a polymorphic region selected from the group consisting of the polymorphic region corresponding to position 957 of SEQ ID NO:6; the polymorphic region corresponding to position 1049 of SEQ ID NO:6; the polymorphic region corresponding to position 1164 of SEQ ID NO:6; the polymorphic region
10 corresponding to position 1526 of SEQ ID NO:6; the polymorphic region corresponding to position 1661 of SEQ ID NO:6; and the polymorphic region corresponding to position 1662 of SEQ ID NO:6.

FIGURE 1

```

1  CTGCAGTGAC  CACTGCCCCA  TCATTGCTGG  CTGAGGTGGT  TGGGGTCCAT  CTGGCTATCT
61  GGGCAGCTGT  TCTCTTCTCT  CCTTTCTCTC  CTGTTTCCAG  ACATGCAGTA  TTTCCAGAGA
121  GAAGGGGCCA  CTCTTTGGCA  AAGAACCTGT  CTAACTTGCT  ATCTATGGCA  GGACCTTTGA
181  AGGGTTCACA  GGAAGCAGCA  CAAATTGATA  CTATTCCACC  AAGCCATCAG  CTCCATCTCA
241  TCCATGCCCT  GTCTCTCCTT  TAGGGGTCCC  CTTGCCAACA  GAATCACAGA  GGACCAGCCT
301  GAAAGTGCAG  AGACAGCAGC  TGAGGCACAG  CCAAGAGCTC  TGGCTGTATT  AATGACCTAA
361  GAAGTCACCA  GAAAGTCAGA  AGGATGCATA  GCAGAGGCCC  AGCAATCTCA  GCTAAGTCAA
421  CTCCACCAGC  CTTTCTAGTT  GCCCACTGTG  TGTACAGCAC  SCTGGTAGGG  ACCAGAGCCA
481  TGACAGGGAA  TAAGACTAGA  CTATGCCCTT  GAGGAGCTCA  CCTCTGTTCA  GGGAAACAGG
541  CGTGGAACA  CAATGGTGGT  AAAGAGGAAA  GAGGACAATA  GGATTGCATG  AAGGGGATGG
601  AAAGTGCCCA  GGGGAGGAAA  TGGTTACATC  TGTGTGAGGA  GTTTGGTGAG  GAAAGACTCT
661  AAGAGAAGGC  TCTGTCTGTC  TGGGTTTGGG  AGGATGTGTA  GGAGTCTTCT  AGGGGGCACA
721  GGCACACTCC  AGGCATAGGT  AAAGATCTGT  AGGTGTGGCT  TGTTGGGATG  AATTTCAAGT
781  ATTTTGGGAA  GAGGACAGCC  ATAGAGACAA  GGGCARGAGA  GAGGCGATTT  AATAGATTTT
841  ATGCCAATGG  CTCCACTTGA  GTTTCTGATA  AGAACCCAGA  ACCCTTGGAC  TCCCAGTAA
901  CATGATTGA  GTTGTTTATG  ATACCTCATA  GAATATGAAC  TCAAAGGAGG  TCAGTGAGTG
961  GTGTGTGTGT  GATTCTTTGC  CAACTTCCAA  GGTGGAGAAG  CCTCTTCCAA  CTGCAGGCAG
1021  AGCACAGGTG  GCCCTGCTAC  TGGCTGCAGC  TCCAGCCCTG  CCTCCTTCTC  TAGCATATAA
1081  ACAATCCAAC  AGCCTCACTG  AATCACTGCT  GTGCAGGGCA  GGAAAGCTCC  ATGCACATAG
1141  CCCAGCAAAG  AGCAACACAG  AGCTGAAAGG  AAGACTCAGA  GGAGAGAGAT  AAGTAAGGAA
1201  AGTAGTGATG  GCTCTCATCC  CAGACTTGGC  CATGGAAACC  TGGCTTCTCC  TGGCTGTCAG
1261  CCTGGTGCTC  CTCTATCTGT  GAGTAACTGT  TCAGGCTCCT  CTTCTCTGTT  TCTTGACTT
1321  GGGGTCGTAA  TCAGGCCTCT  CTTTT

```

FIGURE 2

```

1  GATCTCAGAT  ATCCCTTCTA  TCTACACATT  ATCTATAATT  CTTTCTTTCT  GTAAACTGAA
61 AGGTCCTAGA  AGGAGCCGCA  GCTCAGCAGG  AGAGAGGAGG  AGCTGAGCTG  GGACCCCTAC
121 CTCCTGAGGA  ATGAAATGAT  TATTATAAAG  ACAGCAACCG  AGCTTATTTT  ACCCAAAATA
181 AGGTAGTATA  TTTCTGTTAG  AGTTTAGAGT  TTCATGAGTC  AGGGACCAAG  TTATTGCTTT
241 TCTTTGCCCT  GTATAAAGGC  TTCTCCAAGG  CCTTTGACTT  ACCTAAGTAC  TAAATGTTAT
301 AAAACCAAAC  TCTTCTGACC  TCTCAATCTA  GTCAACTGGG  GCTGTAATTA  TTAATGAAAT
361 TAATGTTTAT  TTTGAAAATA  ATTTACTAGA  CTGAATTACG  AAATCCTGAA  TCATTGTACA
421 CTATCAGTAA  ATATTGGTGG  ACCCAACTGA  ACTGAATGTT  TTGCTTGAAA  TGAAACCTTT
481 GAGATGCAGG  GCTTATGGGT  TCTAGTCCCA  GCTCTAGCAC  TAGCAGACAG  CATGTTCTTG
541 GCTAAGATAC  TGAATCTTCA  AGGCTCAGCT  TCCTCATTCC  GGAAATGGGT  CAATTTTATT
601 GTAAGCAGAG  GTAATTGAGA  GATTCAAAAG  GGACATGAGG  TGTAACAATT  CTCTGTAAAT
661 TGTTAGAATC  CCTGTTAAAA  ATGACCAGTA  AAGCTTTGTG  CAACTGTGTC  TTGACATAAC
721 TTTATTTTTT  TTAATAAAAG  AAATGGAAAT  AACCTCACTA  GGAATTTTAG  AACAAATATG
781 ATGATATCTT  TAAAGAAAAT  GGCTTTGCAC  AAGTATTGAC  ATTAATGATC  TAGTAAAGTG
841 TATCTTTCTA  GTTGTATTTA  GATCCTCAAC  TCAGTATGTC  AGCTCCTGTT  AAGGTCTATA
901 CATTGTGGTG  GTTCTGTGCT  GTGGGTCCAT  TTAGTGATTT  CCCTACCTCC  CATCTTCTAT
961 TGCATCCACA  ACTGTGGTTC  TGTCCATAAT  TTCTTTTGCT  TTCTGTGCAT  TATTACATCA
1021 TATCTGAAAA  TGAGAAACCA  AAAACAATRG  AAAGCAGCCA  TGTCTGGAGG  TGAATTTTAT
1081 CTCGAGAAGC  CCTAGTTTCT  CAAACCTTA  GCACCAAATT  TTTCCCTCAG  TTACATGAG
1141 CGTTTCACTT  CTGCAGTGAT  GGARAAGGGA  GATCCCTTAT  TTCTTCTCAT  GAGCATCTCT
1201 GGTGCTGTTT  CCCTTAGAGA  CAAATAAGGG  GTTCTATTTA  ATGTGAAGCC  TGTTTTATGA
1261 ACAGAATAAA  TGTGGTGTAT  ATTCAGAATA  ACTAATGTTT  GGAAGTTGTT  TTATTTTTCG
1321 TAAAAATTGT  TCTCAAGGCA  GCTCTGGTGT  AAGAGATAAT  ACACCACGAT  GGGCATCAGA
1381 AGACCTCAGC  TCAAATCCCA  GTTCTGCCAG  CTATGAGCTG  TGTGGCACCA  ACAGGTGTCC
1441 TGTTCTCCCA  GGGTCTCCCT  TTTCCCATTT  GAAAAATAAA  AAATAACAAT  TCCTGCCTTC
1501 AGGAATTTTT  TTTAGGGGGT  TTAATKGTAA  AGGTGTTTAT  ATCTGCTAAG  GTAATTTACT
1561 TGATATATGT  TTGGTTATTT  AAGATATATG  AGTTATGTTA  GCTATTTTAT  GTTTAGGCTG
1621 CTGTATTTTT  AGTAGGCTAT  ATTAATATTT  TGAAAGGATT  WMATTATAAA  GAACAAAGTC
1681 TCCTAATCTT  TGATATAGCA  TTGACATACT  TTTTAAATAT  ACAAGGCATA  GAATATGGCC
1741 ATTTCTGTTA  AATCATATAT  TCCCAACTGG  TTATTAATCT  AAGAATTCAG  AATTTTGAGT
1801 AATTGCTTTT  GCATCAGATT  ATTTACTTCA  GTGCTCTCAA  TTATGATGGT  GCATTAGAAC
1861 CATCTGGGTT  AACATTTGTT  TTTTATTACC  AATACCTAGG  CTCCAACCAA  GTACAGTGAA
1921 ACTGGAATGT  ACAGAGTGGA  CAATGGAACG  AAGGAGAACA  AGACCAAAGG  ACATTTTATT
1981 TTTATCTGTA  TCAGTGGGTC  AAAGTCCCTT  CAGAAGGAGC  ATATAGTGGA  CCTAGGTGAT
2041 TGGTCAATTT  ATCCATCAAA  GAGGCACACA  CCGAATTAGC  ATGGAGTGTT  ATAAAAGGCT
2101 TGGAGTGCAA  GCTCATGGTT  GTCTTAACAA  GAAGAGAAGG  CTTCAATGGA  TTCTCTTGTG
2161 GTCCTTGTC  TCTGTCTCTC  ATGTTTGCTT  CTCTTTTCAC  TCTGGAGACA  GAGCTCTGGG
2221 AGAGGAAAAC  TCCCTCCTGG  CCCCACCTCT  CTCCCAGTGA  TTGGAAATAT  CCTACAGATA
2281 GGTATTAAGG  ACATCAGCAA  ATCCTTAACC  AATGTAAGTA  TGCTCCTTCA  GTGGCTTGCA
2341 AAAGGTAAGT  AAATTCACCT  GTATTTTTTT  AATAAAGTGT  ATCCCTAGAG  GTACATGTTA
2401 CAAGAGGTAA  TGGTAAAGTA  AAATACTTTG  AAAGGCTT

```

SEQUENCE LISTING

<110> Gemini Genomics plc

<120> Detection of CYP3A4 and CYP2C9 Polymorphisms

<130> GG119.3PCT

<150> GB 0021286.0

<151> 2000-08-30

<160> 72

<170> PatentIn version 3.1

<210> 1

<211> 1345

<212> DNA

<213> homo sapiens

<400> 1
ctgcagtgac cactgcccc tcatgtgctgg ctgaggtggg tgggggtccat ctggctatct 60
gggcagctgt tctcttctct cttttctctc ctgtttccag acatgcagta tttccagaga 120
gaaggggcca ctctttggca aagaacctgt ctaacttgct atctatggca ggacctttga 180
agggttcaca ggaagcagca caaattgata ctattccacc aagccatcag ctccatctca 240
tccatgccct gtctctcctt taggggtccc cttgccaca gaatcacaga ggaccagcct 300
gaaagtgcag agacagcagc tgaggcacag ccaagagctc tggctgtatt aatgacctaa 360
gaagtcacca gaaagtcaga aggatgcata gcagaggccc agcaatctca gctaagtcaa 420
ctccaccagc ctttctagtt gccactgtg tgtacagcac sctggtaggg accagagcca 480
tgacagggaa taagactaga ctatgccctt gaggagctca cctctgttca gggaaacagg 540
cgtggaaaca caatgggtgg aaagaggaaa gaggacaata ggattgcatg aaggggatgg 600
aaagtgccca ggggaggaaa tggttacatc tgtgtgagga gtttggtgag gaaagactct 660

```

aagagaaggc tctgtctgtc tgggtttgga aggatgtgta ggagtcttct agggggcaca    720
ggcacactcc aggcataagg aaagatctgt aggtgtggct tgttgggatg aatttcaagt    780
attttggaat gaggacagcc atagagacaa gggcargaga gaggcgattt aatagatttt    840
atgccaatgg ctccacttga gtttctgata agaaccacaga acccttggac tccccagtaa    900
cattgattga gttgtttatg atacctcata gaatatgaac tcaaaggagg tcagtgagtg    960
gtgtgtgtgt gattctttgc caacttccaa ggtggagaag cctcttccaa ctgcaggcag   1020
agcacagggt gccctgctac tggctgcagc tccagccctg cctccttctc tagcatataa   1080
acaatccaac agcctcactg aatcactgct gtgcagggca ggaaagctcc atgcacatag   1140
cccagcaaag agcaacacag agctgaaagg aagactcaga ggagagagat aagtaaggaa   1200
agtagtgatg gctctcatcc cagacttggc catggaaacc tggcttctcc tggctgtcag   1260
cctggtgctc ctctatctgt gagtaactgt tcaggctcct cttctctggt tcttggactt   1320
ggggtcgtaa tcaggcctct ctttt                                     1345

```

<210> 2

<211> 19

<212> DNA

<213> synthetic

<400> 2

acaagggcaa gagagaggc

19

<210> 3

<211> 19

<212> DNA

<213> synthetic

<400> 3

acaagggcag gagagaggc

19

<210> 4

<211> 10

<212> DNA

<213> synthetic

<400> 4
agggcaagag 10

<210> 5

<211> 10

<212> DNA

<213> synthetic

<400> 5
agggcaggag 10

<210> 6

<211> 2438

<212> DNA

<213> homo sapiens

<400> 6
gatctcagat atcccttcta tctacacatt atctataatt ctttctttct gtaaaactgaa 60
aggtcctaga aggagccgca gctcagcagg agagaggagg agctgagctg ggacccttac 120
ctcctgagga atgaaatgat tattataaag acagcaaccg agcttatttt acccaaaata 180
aggtagtata tttctgtttag agtttagagt ttcattgagtc agggaccaag ttattgcttt 240
tctttgccct gtataaaggc ttctccaagg ctttgactt acctaagtac taaatgttat 300
aaaaccaaac tcttctgacc tctcaatcta gtcaactggg gctgtaatta ttaatgaaat 360
taatgtttat tttgaaaata atttactaga ctgaattacg aaatcctgaa tcattgtaca 420
ctatcagtaa atattggtgg acccaactga actgaatgtt ttgcttgaaa tgaaaccttt 480
gagatgcagg gcttatgggt tctagtccca gctctagcac tagcagacag catgttcttg 540
gctaagatac tgaatcttca aggtcagct tcctcattcc ggaaatgggt caattttatt 600
gtaagcagag gtaattgaga gattcaaaag ggacatgagg tgtaacaatt ctctgtaaat 660
tgtagaatac cctgttaaaa atgaccagta aagctttgtg caactgtgtc ttgacataac 720
tttatttttc ttaataaaag aaatggaaat aacctacta gggaatttag aacaaatatg 780
atgatatctt taaagaaaat ggctttgcac aagtattgac attaatgatc tagtaaagtg 840
tatctttcta gttgtattta gatcctcaac tcagtatgtc agctcctgtt aaggctcata 900
cattgtggtg gttctgtgct gtgggtccat ttagtgattt ccctacctcc catcttytat 960
tgcattcaca actgtgggtt tgtccataat ttcctttgct ttctgtgcat tattacatca 1020

```

tatctgaaaa tgagaaacca aaaacaatrg aaagcagcca tgtctggagg tgactggggg 1080
gtcgagaagc cctagtttct caaaccctta gcaccaaatt tttccctcag ttacactgag 1140
cgtttcactt ctgcagtgat ggaraagggg gatcccttat ttcttctcat gagcatctct 1200
ggtgctgttt cccttagaga caaataaggg gttctattta atgtgaagcc tgttttatga 1260
acagaataaa tgtggtgtat attcagaata actaatgttt ggaagtgtgt ttatttttgc 1320
taaaaattgt tctcaaggca gctctgggtg aagagataat acaccacgat gggcatcaga 1380
agacctcagc tcaaataccca gttctgccag ctatgagctg tgtggcacca acaggtgtcc 1440
tgttctccca ggggtctccct tttcccattt gaaaaataaa aaataacaat tcctgccttc 1500
aggaattttt tttagggggg ttaatkgtaa aggtgtttat atctgctaag gtaatttact 1560
tgatatatgt ttggttattt aagatatatg agttatgtta gctatttcat gtttaggctg 1620
ctgtattttt agtaggctat attaaatatt tgaaaggatt wmattataaa gaacaaagtc 1680
tcctaatactt tgatatagca ttgacatact ttttaaatac acaaggcata gaatatggcc 1740
atctctgtta aatcatatat tcccaactgg ttattaatct aagaattcag aattttgagt 1800
aattgctttt gcatcagatt atttacttca gtgctctcaa ttatgatggg gcattagaac 1860
catctggggt aacatttggt ttttattacc aatacctagg ctccaaccaa gtacagtga 1920
actggaatgt acagagtgga caatggaacg aaggagaaca agaccaaagg acattttatt 1980
tttatctgta tcagtgggtc aaagtccttt cagaaggagc atatagtgga cctagggtgat 2040
tggtcaattt atccatcaaa gaggcacaca ccgaattagc atggagtgtt ataaaaggct 2100
tggagtgcaa gctcatgggt gtcttaacaa gaagagaagg cttcaatgga ttctcttggt 2160
gtccttggtc tctgtctctc atgtttgctt ctcccttcac tctggagaca gagctctggg 2220
agaggaaaac tccctcctgg cccactcct ctcccagtga ttggaaatat cctacagata 2280
ggtattaagg acatcagcaa atccttaacc aatgtaagta tgctccttca gtggcttgca 2340
aaaggtaagt aaattcacct gtatttttta aataaagtgt atccctagag gtacatgtta 2400
caagaggtaa tggtaaagta aaatactttg aaaggctt 2438

```

<210> 7

<211> 20

<212> DNA

<213> synthetic

<400> 7

ccagcctgaa agtgcagaga

20

<210> 8

<211> 25

<212> DNA

<213> synthetic

<400> 8

tcttagagtc tttcctcacc aaact

25

<210> 9

<211> 20

<212> DNA

<213> synthetic

<400> 9

catgccctgt ctctccttta

20

<210> 10

<211> 19

<212> DNA

<213> synthetic

<400> 10

ccatcccctt catgcaatc

19

<210> 11

<211> 11

<212> DNA

<213> synthetic

<400> 11

agcacctgg t

11

<210> 12

<211> 11

<212> DNA

<213> synthetic

<400> 12
agcacgctgg t 11

<210> 13

<211> 11

<212> DNA

<213> synthetic

<400> 13
accagggtgc t 11

<210> 14

<211> 11

<212> DNA

<213> synthetic

<400> 14
accagcgtgc t 11

<210> 15

<211> 11

<212> DNA

<213> synthetic

<400> 15
gtgtgtacag c 11

<210> 16

<211> 11

<212> DNA

<213> synthetic

<400> 16
gctgtacaca c 11

<210> 17
<211> 11
<212> DNA
<213> synthetic

<400> 17
tggtccctac c 11

<210> 18
<211> 11
<212> DNA
<213> synthetic

<400> 18
ggtagggacc a 11

<210> 19
<211> 25
<212> DNA
<213> synthetic

<400> 19
cactagggaa tttagaacaa atatg 25

<210> 20
<211> 23
<212> DNA
<213> synthetic

<400> 20
gcacagaaag caaaggaaat tat 23

<210> 21
<211> 27
<212> DNA
<213> synthetic

<400> 21
tgtatttaga tcctcaactc agtatgt 27

<210> 22
<211> 21
<212> DNA
<213> synthetic

<400> 22
ggatctccct tctccatcac t 21

<210> 23
<211> 23
<212> DNA
<213> synthetic

<400> 23
ggtccattta gtgatttccc tac 23

<210> 24
<211> 25
<212> DNA
<213> synthetic

<400> 24
atacaccaca tttattctgt tcata 25

<210> 25
<211> 22
<212> DNA
<213> synthetic

<400> 25
ccaaattttt ccctcagtta ca 22

<210> 26
<211> 20
<212> DNA
<213> synthetic

<400> 26
ttggtgccac acagtcata 20

<210> 27
<211> 20
<212> DNA
<213> synthetic

<400> 27
gccttcagga atttttttta 20

<210> 28
<211> 25
<212> DNA
<213> synthetic

<400> 28
ccagttggga atatatgatt taaca 25

<210> 29
<211> 25
<212> DNA
<213> synthetic

<400> 29
gctgctgtat ttttagtagg ctata 25

<210> 30
<211> 22
<212> DNA
<213> synthetic

<400> 30
cgttccattg tccactctgt ac 22

<210> 31

<211> 20

<212> DNA

<213> synthetic

<400> 31
tcaaggcagc tctggtgtaa 20

<210> 32

<211> 25

<212> DNA

<213> synthetic

<400> 32
agttgggaat atatgattta acaga 25

<210> 33

<211> 11

<212> DNA

<213> synthetic

<400> 33
atcttctatt g 11

<210> 34

<211> 11

<212> DNA

<213> synthetic

<400> 34
atcttttatt g 11

<210> 35

<211> 11

<212> DNA

<213> synthetic

<400> 35

acaatagaaa g

11

<210> 36

<211> 11

<212> DNA

<213> synthetic

<400> 36

acaatggaaa g

11

<210> 37

<211> 11

<212> DNA

<213> synthetic

<400> 37

atggagaagg g

11

<210> 38

<211> 11

<212> DNA

<213> synthetic

<400> 38

atggaaaagg g

11

<210> 39

<211> 11

<212> DNA

<213> synthetic

<400> 39
ttaatggtaa a 11

<210> 40

<211> 11

<212> DNA

<213> synthetic

<400> 40
ttaattgtaa a 11

<210> 41

<211> 12

<212> DNA

<213> synthetic

<400> 41
ggatttcatt at 12

<210> 42

<211> 12

<212> DNA

<213> synthetic

<400> 42
ggattaaatt at 12

<210> 43

<211> 11

<212> DNA

<213> synthetic

<400> 43
caatagaaga t 11

<210> 44
<211> 11
<212> DNA
<213> synthetic

<400> 44
caataaaaga t 11

<210> 45
<211> 11
<212> DNA
<213> synthetic

<400> 45
ctttctattg t 11

<210> 46
<211> 11
<212> DNA
<213> synthetic

<400> 46
ctttccattg t 11

<210> 47
<211> 11
<212> DNA
<213> synthetic

<400> 47
cccttctcca t 11

<210> 48
<211> 11
<212> DNA
<213> synthetic

<400> 48
cccttttcca t 11

<210> 49

<211> 11

<212> DNA

<213> synthetic

<400> 49
tttaccatta a 11

<210> 50

<211> 11

<212> DNA

<213> synthetic

<400> 50
tttacaatta a 11

<210> 51

<211> 12

<212> DNA

<213> synthetic

<400> 51
ataatgaaat cc 12

<210> 52

<211> 12

<212> DNA

<213> synthetic

<400> 52
ataatttaat cc 12

<210> 53
<211> 11
<212> DNA
<213> synthetic

<400> 53
tacctcccat c 11

<210> 54
<211> 11
<212> DNA
<213> synthetic

<400> 54
aaccaaaaac a 11

<210> 55
<211> 11
<212> DNA
<213> synthetic

<400> 55
ctgcagtgat g 11

<210> 56
<211> 11
<212> DNA
<213> synthetic

<400> 56
taggggggttt a 11

<210> 57
<211> 11
<212> DNA
<213> synthetic

<400> 57
atttgaaagg a 11

<210> 58

<211> 11

<212> DNA

<213> synthetic

<400> 58
gatgggaggt a 11

<210> 59

<211> 11

<212> DNA

<213> synthetic

<400> 59
tgtttttggt t 11

<210> 60

<211> 11

<212> DNA

<213> synthetic

<400> 60
catcactgca g 11

<210> 61

<211> 11

<212> DNA

<213> synthetic

<400> 61
taaaccct a 11

<210> 62
<211> 11
<212> DNA
<213> synthetic

<400> 62
tcctttcaaa t 11

<210> 63
<211> 11
<212> DNA
<213> synthetic

<400> 63
tgtggatgca a 11

<210> 64
<211> 11
<212> DNA
<213> synthetic

<400> 64
catggctgct t 11

<210> 65
<211> 11
<212> DNA
<213> synthetic

<400> 65
agggatctcc c 11

<210> 66
<211> 11
<212> DNA
<213> synthetic

<400> 66
taaacacctt t 11

<210> 67

<211> 11

<212> DNA

<213> synthetic

<400> 67
tggtctttat a 11

<210> 68

<211> 11

<212> DNA

<213> synthetic

<400> 68
ttgcatccac a 11

<210> 69

<211> 11

<212> DNA

<213> synthetic

<400> 69
aagcagccat g 11

<210> 70

<211> 11

<212> DNA

<213> synthetic

<400> 70
gggagatccc t 11

<210> 71

<211> 11

<212> DNA

<213> synthetic

<400> 71

aaaggtgttt a

11

<210> 72

<211> 11

<212> DNA

<213> synthetic

<400> 72

tataaagaac a

11